Coarticulatory effects in perception

Santiago Fernández, Sergio Feijóo

Departamento de Física Aplicada
Universidad de Santiago, 15782 Santiago, Spain
fasergio@usc.es

Abstract

The perceptual interaction between adjacent CV segments is studied in Fricative-Vowel syllables from a coarticulatory point of view. The results of a perceptual experiment with conflicting-cue stimuli are analyzed under the assumption, supported by the results, of the possible influence of F-to-V carryover coarticulation on the integration process. The DAC scale was used to estimate the degree of F-to-V carryover coarticulation in the original syllables. The magnitude of the coarticulatory effect permitted us to derive a prediction of the perceptual results based on the articulatory compatibility between the fricative and the vowel. The correlation between actual and predicted decrease in perceptual identification caused by the insertion of a conflicting transition was computed. The results show that the coarticulatory processes cannot explain the outcome of the perceptual experiment. Nevertheless, the perceptual role played by the /i/ transition can be effectively explained as a consequence of the F-to-V coarticulation.

1. Introduction

Phonetic integration in CV syllables is a yet unexplained process. Perceptual identification of the consonant is better when listeners are presented with both the consonant and the vowel with respect to the presentation of only the consonant. A possible explanation for that effect is based on the articulatory compatibility between the place of articulation of the consonant and the vowel: the articulatory constraints imposed by the consonant on the vowel would make it necessary to have both parts to properly evaluate consonant identity.

To study this issue a perceptual experiment with conflicting-cue stimuli was carried out with fricatives /f, θ, s, ʃ/ in the context of vowels /a,e,i,o,u/. In that experiment the vocalic transition of a particular FV pair was removed, and the vocalic transition corresponding to the same vowel in another consonantal context, or to a different vowel in the same consonantal context, is appended to the fricative noise. The effect of trading the original transition against a conflicting transition on the articulatory compatibility between consonant and vowel can be estimated using the DAC scale proposed by [1]. If the coarticulatory effects are responsible for the auditory effects, the decrease in perceptual identification of the fricatives when the proper transition is traded against a conflicting transition, should be correlated with the predictions based on the articulatory compatibility between the consonant and the vowel.

2. Articulatory compatibility

The core of the DAC scale is the degree of articulatory constraint imposed on the production of the phonemes, particularly by the use of the tongue body to form the necessary constriction. For instance, palatalized consonants and vowels are more constrained than labials, since tongue dorsum raising is necessary for their production, while in labials the tongue is reasonably free to adopt a configuration compatible with the following phoneme. For instance, in a syllable like /ʃa/, the fricative is highly constrained, while the vowel is comparatively less constrained. Little V-to-F anticipatory coarticulation is expected (coarticulatory influence of the vowel on the fricative). On the other hand, some F-to-V carryover coarticulation is expected (coarticulatory influence of the consonant on the vowel), since the fricative is more constrained than the vowel. In a syllable like /fi/, for instance, the fricative is unconstrained since the tongue is not used to form the constriction. Then, little F-to-V carryover coarticulation is expected. Since the vowel is more constrained than the fricative, some V-to-F anticipatory coarticulation is expected.

3. Perception of conflicting-cue stimuli

In a previous paper, results of the perception of conflicting-cue stimuli are reported [2]. One man and one woman pronounced FV syllables made up of the combination of fricatives /f, θ, s, ʃ/ with vowels /a,e,i,o,u/. 51 listeners were asked to identify the initial fricative in the natural and conflicting-cue stimuli.It was found that in FV syllables, F-to-V carryover coarticulation could play some role in the perception of the fricatives, while V-to-F anticipatory coarticulation caused little or no effect. Then, in the present paper only the effect of trading the proper vocalic transition against a conflicting transition of the same vowel in another consonantal context is considered.

The stimuli that will be analyzed were constructed in the following way. Let us consider a natural FV syllable denoted by $F^V V^F$. If the vowel V^F is removed and the same vowel from another fricative context $V^{F'}$ is appended to the original fricative noise, then we have the conflicting-cue stimuli $F^V V^{F'}$. That kind of stimuli allows us to study the influence of the F-to-V carryover coarticulation. Another possibility is to remove the vowel V^F and append another vowel from the same fricative context V'^F, obtaining the conflicting-cue stimulus $F^V V'^F$. That kind of stimuli allows us to study the influence of the V-to-F anticipatory coarticulation. The results of that perceptual experiment showed that the perceptual interaction between the consonant and the vowel in fricative-vowel syllables cannot be explained only by the coarticulatory influence of the consonant or the vowel on adjacent segments. The effects of the F-to-V carryover coarticulation are perceptually irrelevant for /ʃ/. For /s/ those effects were more important, especially in the context of /a/. For /θ/ those effects were important in the /a/, /o/, and /u/

contexts, and irrelevant in the /i/, /e/ contexts. For /f/, the perceptual effects of coarticulation are important in the /a/, /e/, /o/, and /u/ contexts, and irrelevant in the /i/ context. On the other hand, the effects of the V-to-F anticiaptory coarticulation were perceptually irrelevant. Overall, the results suggest that the link between coarticulatory effects and perception is quite weak, at least for some fricatives.

How can we estimate the perceptual effect of changing the vocalic transition? Let us denote by $P(F^V V^F)$ the number or percent of listeners who correctly identified fricative F in stimulus $F^V V^F$. Let us denote by $P(F^V V^{F'})$ the number or percent of listeners who correctly identified fricative F in stimulus $P(F^V V^{F'})$. Then, the decrease in identification caused by the insertion of the conflicting transition can be tentatively estimated as

- $DCC = 0$, if $P(F^V V^F) < P(F^V V^{F'})$
- $DCC = P(F^V V^F) - P(F^V V^{F'})$, otherwise

It is assumed that identification of the natural syllable is at least as good or better than the identification of the conflicting-cue syllable. Then, DCC should be null whenever there is a certain increase (instead of decrease) in the number of listeners that identify the proper consonant in the syllable. This case happened a few times, but, as assumed, the magnitude of the improvement was very small.

4. Predictions based on the DAC scale

If phonetic integration is a side effect of the articulatory processes that take place in the production of speech, then it would be possible to predict the perceptual effects caused by the conflicting transitions by considering the articulatory compatibility between the fricative and the vowel.

The following values can be tentatively assigned to the phonemes involved ([1]):

- $DAC = 1$: /f/
- $DAC = 1.5$: /a/, /o/, /u/
- $DAC = 2$: /θ/
- $DAC = 2.5$: /s/, /e/
- $DAC = 3$: /ʃ/, /i/

The DAC scale proceeds from less constrained ($DAC = 1$) to more constrained phonemes ($DAC = 3$). Of course, this assignment of DAC values is only one of several possibilities. It is not clear yet which values should be assigned to some phonemes, particularly those with intermediate values. While it is easy to agree on the phonemes with more extreme values (like /f/, /i/ or /ʃ), there is simply no objective procedure to tell whether /a/ should have the same DAC value as /o/, or whether that value should be 1.5, 1.25 or 2. On the other hand, the DAC scale considered is intimately related to the role played by tongue dorsum, and does not consider other effects like lip rounding. As a result, the proposed DAC values bear a relation to F2 values [1]. Consonantal effects reflect trends in lingual and jaw coarticulation, whereas vocalic effects are related to jaw coarticulation [4].

Using the proposed DAC scale, it is possible to estimate the degree of F-to-V carryover coarticulation for a particular FV syllable. DAC values should be related to the extent to which a certain segment will influence adjacent segments, considering also that in order to evaluate the magnitude of the coarticulatory effect, we have to consider the DAC value of the contextual segment and whether the constraints involved are compatible or opposing.

4.1. Degree of F-to-V carryover coarticulation

We shall denote by DFV the degree of influence of the fricative on the following vowel. The DFV values can be computed as follows :

- $DFV = 0$, if $DAC(V) \geq DAC(F)$
- $DFV = DAC(F) - DAC(V)$, otherwise

The degree of F-to-V carryover coarticulation is minimal (0 valued) whenever the vowel is more constrained than the consonant, or equally constrained. The degree of coarticulation in other cases is obtained as the difference in DAC value between consonant and vowel, whenever the consonant is more constrained than the vowel. As that difference becomes bigger, it is assumed that the degree of influence of the consonant on the vowel would be bigger too.

The resulting values of DFV would be:

- $DFV = 0$: /fa/, /fe/, /fi/, /fo/, /fu/, /si/, /θi/, /ʃi/, /se/, /θe/
- $DFV = 0.5$: /θa/,/θo/, /θu/,/ʃe/
- $DFV = 1$: /sa/, /so/, /su/
- $DFV = 1.5$: /ʃa/, /ʃo/,/ʃu/

If the perceptual effects are a consequence of the articulatory compatibility between the fricative and the vowel, the DFV values can be used to estimate the decrease in identification caused by a conflicting transition, in the following way. Let us suppose that we have the syllable /θa/. Now we construct the conflicting-cue stimulus /θa^f/, where /a^f/ indicates that the new vowel was excised from syllable /fa/. DFV(/θa/) is 0.5; DFV(/fa/) is 0; If we remove /a^θ/, and instead we place /a^f/, we have the stimulus /θa^f/, and the incompatibility between the new vowel and the original fricative noise can be estimated as 0.5. Let us consider the opposite process. If /a^f/ is removed from /fa/, and instead we place /a^θ/, the resulting syllable, /fa^θ/, would have the same degree of incompatibility between the fricative and the vowel, 0.5, as /θa^f/. If, for instance, a vowel is removed from a FV syllable with $DFV = 0$, and is replaced by the vowel from a syllable with $DFV = 0$, then no effect would be expected. That change should not cause any decrease in the perceptual identification of the fricative in the new syllable.

Then, the perceptual effect can be predicted by

- $DCC' = DFV(F^V V^F) - DFV(F^V V^{F'})$ for pair /θ/ and /s/
- $DCC' = DFV(F^V V^F) + DFV(F^V V^{F'})$ for any other pair

The fact of considering the specific case of the pair /θ/ and /s/ is due to their proximity in place of articulation. Moreover, the results of the conflicting-cue experiments confirmed that the perceptual effects of the vocalic transitions of both fricatives are quite similar. Let us assume that, for instance, we have syllable /θa/, and the vowel is replaced by /a^s/. The DFV value of /sa/ is 1, while that of /θa/ is 0.5. If it is assumed that both consonants have a compatible place of articulation, then the effect of trading the vowel cannot be estimated as the sum of both DFV values. As /a^s/ has a stronger F-to-V coarticulation than /a^θ/, the incompatibility between the new vowel and the original fricative noise in the new syllable /θa^s/ , can be estimated as the difference between both DFV values, which is 0.5, rather than 1.5.

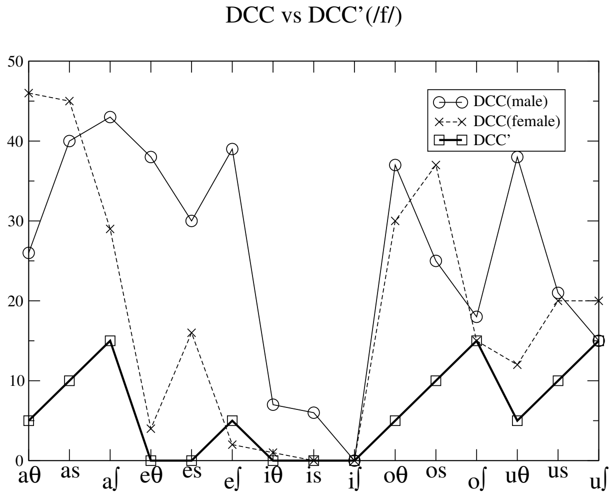

Figure 1: *Actual (DCC) and predicted (DCC') decrease in identification caused by the conflicting transition for fricative /f/.*

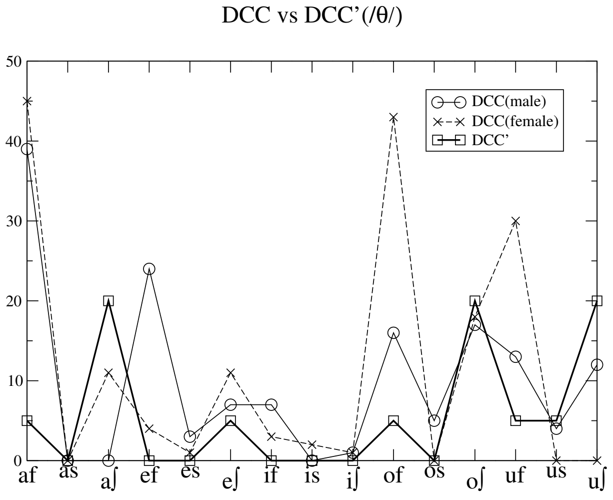

Figure 2: *Actual (DCC) and predicted (DCC') decrease in identification caused by the conflicting transition for fricative /θ/.*

5. Correlation between predicted and actual decrease in identification

Figures 1-4 show the actual (DCC) and predicted (DCC') decrease in identification caused by the conflicting transition, for each fricative separately. DCC' values were rescaled to ten times the original values, due to the differences in magnitude range between DCC and DCC'. The tendency in all four figures is approximately the same: the agreement between DCC and DCC' is occasional, while there seems to be some particular effects reflected in DCC' that are very similar for both speakers. An exception to that rule is represented by the results for the /i/ transition: there is clearly a good agreement between both speakers in DCC', and also a good agreement between DCC and DCC'.

Of course the DAC values assigned to the different phonemes are only a tentative approach. There is simply no way of establishing a completely realistic scale. Nevertheless, it is also clear that they offer a rough description of the expected coarticulatory effects. It should be remembered that we are testing the hypothesis of the possible coarticulatory influence between neighboring segments on the phonetic integration process. If that influence exists, DCC' should predict with a certain degree of accuracy the DCC values obtained in the perceptual experiment.

Table 1: *Correlations between actual (DCC) and predicted (DCC') decrease in identification for the male and female speakers, separately for each fricative*

fricative	male	female
ɵ	0.11	0.16
f	0.24	0.52
s	0.11	0.53
ʃ	-0.19	0.47

DCC and DCC' were correlated. Table 1 shows the results of the correlations for each fricative separately. None of the correlations was significant at the $p < 0.01$ level. For the male speaker, overall correlation between DCC and DCC'considering all the fricatives together was 0.08, while for the female speaker it was 0.27. The overall correlations were non-significant at the 0.01 level. Clearly, there is very little relation between the actual effects caused by the conflicting transitions and the values predicted by the DAC scale. In order to have a reference value for those figures, the DCC values of the male and female speakers were correlated. A significant correlation ($p < 0.0001$) of 0.63 was obtained. What this last figure shows is that the perceptual decrease caused by the conflicting transition is more consistent than the predicitons based on the effect of the F-to-V carryover coarticulation.

6. Influence of coarticulatory effects in perception

One of the problems of the proposed approach is that the values assigned by the DAC scale to a given phoneme are fixed, predicting always the same figure for a particular coarticulatory effect regardless of personal characteristics (like speaking style or anatomical characteristics), that may provide a certain amount of articulatory variability. An additional problem is that the predicted effect is always symmetrical, while the results of the conflicting-cue experiment did not support that hypothesis. For instance, the decrease in identification caused by stimulus /fa^{s}/ is 26 for the male, and 46 for the female speaker, while /sa^{f}/ causes a decrease of 0 and 28, respectively.

This is not to say that the DAC scale is not useful from an articulatory point of view. In fact, a clear relationship between DFV and F2 excursion size values was found in another study [3], indicating that there is a close link between production characteristics and acoustical output. From a perceptual point of view, the coarticulatory processes represented in the DAC scale offer little information.

However, the DAC scale does provide a satisfactory explanation of some of the perceptual effects encountered. For instance, since vowel /i/ is highly constrained, no F-to-V coarticulation is expected for that vowel. As a result, the consonantal context is of little importance for /i/, and trading the /i/ transition from one consonantal context to another should cause little or no decrease in identification. The conflicting-cue ex-

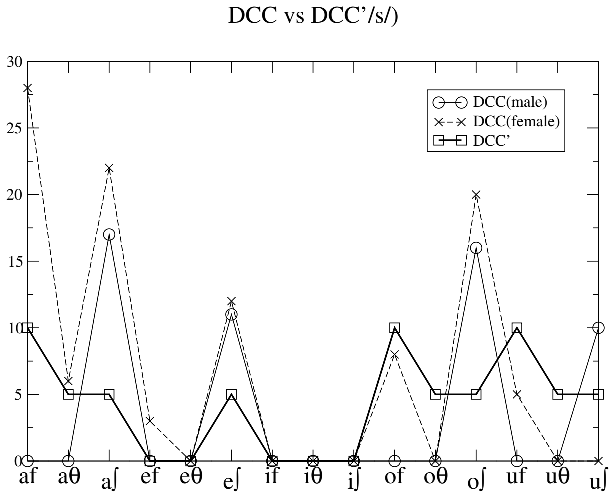

Figure 3: *Actual (DCC) and predicted (DCC') decrease in identification caused by the conflicting transition for fricative /s/.*

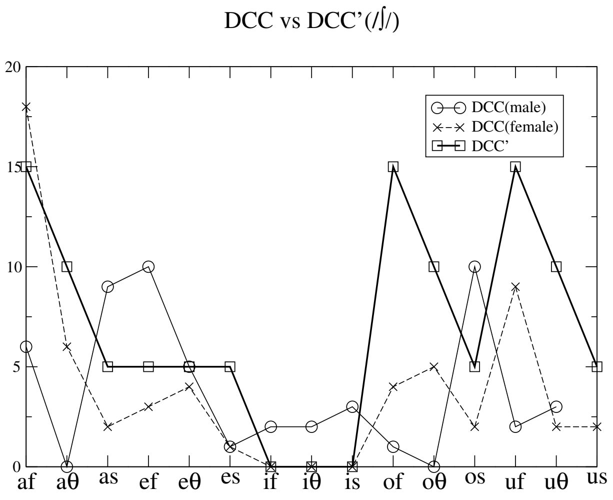

Figure 4: *Actual (DCC) and predicted (DCC') decrease in identification caused by the conflicting transition for fricative /ʃ/.*

periment confirms that this is indeed the case. Nevertheless, in some other instances, the coarticulatory scale fails to predict the perceptual outcome. For instance, F-to-V carryover coarticulation is supposed to be moderate for /su/ and minimal for /fu/. Then, stimulus /su^f/ should be associated with an important decrease in identification, but the perceptual data showed that the decrease was 0 for the male, and 5 for the female speaker.

Table 2: *Actual(DCC) and predicted (DCC') decrease in identification caused by the insertion of the conflicting transition. Values are pooled over each fricative*

fricative	DCC(male)	DCC(female)	DCC'
θ	148	169	8.5
f	383	277	9.5
s	54	104	6.5
ʃ	64	68	10.5

Table 2 shows DCC and DCC' values pooled over each particular fricative. The maximum magnitude effects should be expected for /ʃ/, and the minimum effects should correspond to /s/, as the DCC' values in Table 2 clearly show. The DCC values indicate otherwise. In fact, the asymmetrical behaviour of the listeners in the perceptual experiment suggest that /f/ transitions are well tolerated by alveolar /s/ and palatal /ʃ/, possibly because the fricative noise tends to override the effect of the transitions. On the other hand, palatal, alveolar and dental transitions tend to cause a much bigger impact on the labial fricative, possibly because the noise is a weak cue in this case. Dental, alveolar and palatal transtions seem to be somewhat compatible among them, and trading one against the others does not cause big changes in perception, despite the fact that production requirements vary a great deal among them. For instance, stimulus /θa^ʃ/ causes a decrease in identification of 0 and 11 for the male and the female speakers, respectively, when the carryover coarticulatory influence of /ʃ/ on /a/ is maximum, while that of /θ/ on /a/ is minimum.

7. Conclusions

Overall, the results obtained show that F-to-V carryover coarticulation cannot predict the outcome of a perceptual experiment with conflicting-cue stimuli, originally formed by Fricative-Vowel syllables. The only perceptual effect that seemed to be based on the coarticulatory influence of the fricative on the vowel, is the role played by the /i/ transition: Since /i/ is maximally constrained, trading its transition between fricative contexts causes little or no effect at all. As regards the vocalic transition of other vowels, coarticulatory effects very unlikely provide the cues supposedly used by listeners to integrate information from both the fricative noise and the vocalic transition, in order to identify the fricative.

8. Acknowledgments

This work was financed by Xunta de Galicia under project PGIDT00PXI20608PR. The authors would like to thank D. Recassens for helpful comments on the DAC scale.

9. References

[1] Recassens, D., Pallares, M. D., Fontdevila, J.,"A model of lingual coarticulation based on articulatory constraints", J. of the Acoust. Soc. Amer., 102(1):544–561, 1997.

[2] Fernández, S., Feijóo, S., Balsa, R., Barros, N., "Perceptual effects of coarticulation in fricatives", Proc. of the ICASSP 2000, 2000

[3] Feijóo, S. and Fernández, S., "Acoustical effects of coarticulation in Speech", International Congress on Acoustics (Rome), 2001.

[4] Recassens, D., Pallares, M. D.,"A study of F1 coarticulation in VCV sequences", J. of Speech, Lang. and Hear. Res., 43:501–512, 2000.